

AD-A074 654

WEST VIRGINIA UNIV MORGANTOWN

F/G 5/11

AN EXPERIMENT ON THE DEGRADATION OF ACCURACY IN HUMAN RECALL OF--ETC(U)

JUL 79 H R BERNARD, P D KILLWORTH, L SAILER

N00014-75-C-0441

UNCLASSIFIED BK-118-79

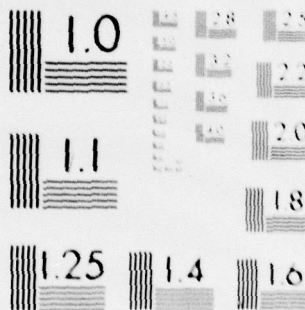
NL

| OF |

AD
A074654



END
DATE
FILMED
10-79
DDC



MICROCOPY RESOLUTION TEST CHART

A074654

DDC FILE COPY

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

RESEARCH AND DEVELOPMENT
EXPERIMENTAL FORM

1. REPORT NUMBER BK-118-79		2. GOVT ACCESSION NO.		3. RECIPIENT CATALOG NUMBER	
4. TITLE (and Subtitle) An experiment on the degradation of accuracy in human recall of communications.				5. TYPE OF REPORT & PERIOD COVERED Interim Dec 78 - Apr 79	
7. AUTHOR(s) 10 H. Russell/Bernard, Peter D./Killworth, and Lee/Sailer				8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0441-P00001	
9. PERFORMING ORGANIZATION NAME AND ADDRESS H. Russell Bernard, Dept. of Anthropology, University of Florida, Gainesville, FL 32611				10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 12 24	
11. CONTROLLING OFFICE NAME AND ADDRESS ONR-Code 452 Arlington, Va 22217				12. REPORT DATE 11 31 Jul 1979	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) West Virginia University Morgantown, WV 26505				13. NUMBER OF PAGES	
				15. SECURITY CLASS. (of this report) Unclassified	
				15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release, distribution unlimited					
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)					
18. SUPPLEMENTARY NOTES					
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) social networks, cognition and behavior, sociometrics					
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) This report describes progress, as of 6/15/79, of an experiment which was conducted on EIES, the Electronic Information Exchange System. The report describes the experiment, the data collection, and the codebooks for the data files constructed so far. The tape of data files is available for secondary analysis.					

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 68 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

79 10 02 025

I. Introduction

This report describes progress, as of June 15, 1979 in what is known as "Russ's experiment" on EIES, the Electronic Information Exchange System. The report describes the experiment, the data collection, and codebooks for the two data files constructed so far.

Social network data are usually collected by asking people to answer some form of the question "who do you _____?" where the blank is filled in by "talk to," "seek advice from," etc. In a recent series of papers (Killworth and Bernard 1976, 1979; Bernard and Killworth 1977, 1979) we have examined the accuracy, or validity of such data. The general approach was to ask the members of some group "Who do you _____?" (in various ways) and compare the answers to data obtained from direct observation of the group. We call these the 'recall' and 'behavior' data sets. Our main conclusion is that, apparently, informants cannot recall with acceptable accuracy who they communicate with in a group over a period of time. For example, informants claim they talk to people they never actually talk to; they claim they never talk to people they do talk to; and they are unable to rank or scale their communications accurately, even when referring to the people with whom they have communicated the most. This means that social network (or sociometric) data, as they are usually collected, can not be used with any reliability as a substitute for the behavior which they supposedly represent.

We considered the possibility that individual differences among informants (on socioeconomic indicators, or on how accurate they felt they were, for example) might help to account for variation in their accuracy. Thus far, five different data sets have been examined, representing a variety of study populations. We have found nothing that accounts for variation in informant accuracy. We also considered the possibility that different structures of groups of communicants might be related to accuracy of communication recall. We tested many different triadic structures, and again found nothing to account for variation in informant accuracy, though we did find that both recall data and actual communication data possess significantly high or low amounts of structure on every structural indicator we could think of. Unfortunately, the structures in any particular set of recall data were never produced by the same triads as those in the matched set of behavior data.

Finally we considered the possibility that informant accuracy is a function of sub-group organization. Perhaps modern clique-finding algorithms might uncover an essential, underlying agreement between recall and behavior data? Again, this turned out not to be the case. The three clique-finders we tested (chosen because they represent three major traditions in the literature) failed to produce similar cliques in our matched sets of recall and behavior data (or with each other).

Of course, it is possible that informant characteristics really are responsible for variations in accuracy of communications recall data (or any behavioral recall data). It may be that we have simply not made the correct comparisons. Similarly, there may be triadic structures which would give better answers than those we have tested; and there are certainly many clique-finders which we have not examined.

In this experiment we examine the possibility that the inaccuracy we have found is a function of time period over which informants are asked to recall their behavior. All our previous data sets have been based on informant recall

of their behavior during one of three "windows": the previous five days; the previous month; and the forthcoming month. Any period of time, or window, can be characterized by two quantities, which we call "lag" and "width." Width is the amount of time over which informants are asked to recall their behavior. Lag is the amount of time that has elapsed since the end of the window. Thus, the five-day windows in some of our previous experiments have a width of five days, and a lag of, at most, one day.

The majority of questions asked by students of social networks have a lag of less than one day, with widths that range from a few days to the life time of the informant. It seems plausible that very recent time windows should tend to be more accurate than windows far in the past. "Who did you talk to one minute ago?" should yield more accurate data than "who did you talk to for a minute at this time last month?" Similar variations in accuracy could be caused by different widths: "who did you talk to during a period of a week, a month ago?" The question addressed in this experiment is "what is the combination of lag and width which yields the most accurate social network data?" A direct answer to this question will be offered in the next paper in this series. In the present paper, we describe the experiments and the data it yielded.

II. The Experiment

During the past three years, under National Science Foundation auspices, the New Jersey Institute of Technology has developed a computer-based conference system known as EIES (Electronic Information Exchange System). A complete description of EIES, including its technology and design philosophy may be found in Turoff and Hiltz (1978). Briefly, EIES allows an individual to exchange messages with others on the system by leaving the message in a central computer for pick-up during the next time the "receiver" logs on. Messages may be addressed to single individuals, with or without copies to other individuals. Messages may also be sent to "groups." A typical group on EIES consists of between 10 and 100 people who have common interests and who are working on a common problem. Many groups on EIES are composed of scientists who hold ongoing "conferences" for periods up to two years (limited only by the young age of EIES). Members of a group are free to enter into small or large conferences with subsets of their own groups, or of other groups.

"Conference comments" are a kind of public message submitted by a conferee for all members of a conference to read. Conference topics range from broad, theoretical discussions of, for example, General Systems Theory, to very specific, work-group discussions of, for example, data manipulation techniques. One EIES group planned and executed the experiment reported in this paper.

"Private messages" are communications between individuals; only the sender or the addressees of a private message are privileged to access that message. Private messages include side remarks about conferences; personal letters between friends, enemies and colleagues; and chit-chat between casual EIES acquaintances. Every EIES participant can be identified and address by name, nickname, or number (e.g., H. RUSSELL BERNARD, RUSS, or 357).

In other words, conferences function like the formal organizations of a business or university department. The private messages replace what might be called the "day-to-day communication network," where people talk about work and more casual social relations. Many studies of social networks in such environments have

been conducted; the advantage of EIES for our purposes is that every non-formal communication (ie., private message) can be permanently recorded.¹

Between December, 1978 and April, 1979, 57 paid volunteer EIES users participated in our experiment. An invitation to participate in the experiment was sent to over 150 EIES members via a personal message from Bernard (see Appendix I). Depending on the rate of their EIES use, each informant took up to 37 interviews, each for a specific lag and width. The informant was given a window and was then asked to list the people with whom he or she communicated during that window. Next, informants were given an opportunity to add or to delete names from the list, and were asked to estimate the number of messages and the number of lines sent to and received from each communicant recalled. Finally, they were asked to rate their confidence, on a scale from 1-7, about the information provided. At the end of each interview, informants were given the opportunity to send the experimenters a message containing any observations or suggestions they wished to make. Twenty-seven windows were established according to the pattern shown for the first 27 windows in Appendix II. Windows were selected for informants in random order. The remaining 10 windows we call "last on;" for these windows people were asked to recall their communications during the last time they were on EIES. This ranged from several weeks to several minutes in lag, and from several minutes to several hours in width.

Two questionnaires were also administered. The first interview collected data on all our informants' age, sex, self-reported EIES use, and self-reported estimates of memory ("how well, on a scale from 1-7, do you remember birthdays?"). The second interview was taken by the 22 informants who completed all 27 of the basic window interviews. It again asked for information on EIES use, and also asked informants to report the 20 people with whom they believed they communicated most. For each of those 20, informants were asked to rate (on a scale of 1-7) the importance of the communication, how satisfying it was, how desirable communication was with that person, and how interesting it was.

Since data collection in this experiment was, in a sense, scheduled at the leisure of the informant, and performed by the central computer itself, it was possible to allow our respondents some control over the progress of interviews. An informant could withdraw from the experiment (permanently or temporarily) at any time. Also, informants could check on their own accuracy for the previously completed interviews by using a routine called "feedback" (see Appendix III). And they could check on their general progress by examining a routine called "windows" (see Appendix II).

Two other routines were introduced which we felt might illuminate the causes of variation in informant accuracy. These were called "raincheck" and the "harassment limit." The interviews were administered randomly at the very beginning of an EIES session at a rate sufficient to keep all the subjects at the same pace. For any given interview, a respondent was allowed to take a raincheck of from 1-7 days. (This was changed to 1-3 days later in the experiment, since we felt things were going too slowly.) After taking a raincheck, there was no way a respondent could avoid an interview the next time he or she logged onto EIES.

The harassment limit was the maximum amount of bother that an informant was willing to put up with in one session. After each interview, which averaged about 6 to 8 minutes, if sufficient time was left in the harassment limit, a "last on" window was administered. Most informants selected 20 minutes as their harassment limit.

III. The Informants

The informants ranged in age from 18 to 64, and the mean amount of EIES experience was about 300 hours. Twenty-three completed all 37 windows and both interviews, and, out of 57 informants, no regular window was taken fewer than 32 times or more than 38 times. Twenty-two informants took all 10 last on windows, and 37 people took at least 1.

On EIES there is a phenomenon called "deleted" messages -- messages sent, and possibly received, but then purged from EIES before our data collection routines could collect them. Eight percent of the 1211 interviews are contaminated by deleted messages, but never by more than 1 message per interview. In general, between 0 and 48 communicants were recalled (mean = 2.46) and between 0 and 111 people was actually communicated with (mean = 6.05).

IV. Measuring Accuracy

Some of the factors that affect the accuracy of recall can be accounted for by measuring accuracy in different ways. For example, people might be better able to recall the messages they send to others than the ones they receive. Or they might recall the people they communicate with, but not the exact (or even approximate) number of messages.

We have concocted 48 different measures of accuracy, most of which were used previously in this series of papers. They fall into broad classes which make them easy to describe.

Each measure is computed separately for messages the informant recalls sending to people, those from people, and those both to and from, combined (shown in the table as $\begin{matrix} T \\ F \\ B \end{matrix}$). Table 1 shows the accuracy measures in concise form.

The first six classes use only the names of those recalled and those actually communicated with. (Measures that use "number of messages," and "number of lines" as indicators of intensity of messaging follow.) T1, T1P, and T2P are straightforward. T12A counts the number of mistakes (T1 + T2) as meaningful in relation to the total number of people actually communicated with. T12AR counts the number of mistakes as a percentage of the total number of possible mistakes (NA + NR), given the number of people recalled and the number of people actually communicated with for that informant and window.

The second and third classes of inaccuracy measures use either "number of messages" or "number of lines" as indicators of intensity of communication, noted in the table as $\begin{matrix} M \\ L \end{matrix}$. This allows us to rank the recalled and actual communicants, and to see, for instance, whether people can recall with accuracy those people with whom they communicate most.

TOP5, TOP3, and TOP1 measure the percentage of errors people make about those they report as their most frequent communicants. WIN2 suggests that people might be able to recall those people most frequently communicated with, but that the exact ranks might be off by 2 or so, and still be counted as correct. WIN10 should

indicate when a person recalls actually communicated with in the correct order, but does not penalize the informant for leaving people out randomly.

So, for example, TIPF is the percentage of messages from others recalled by the informant which in fact did not exist. And TOP5TL is the percentage of people reported to be in the top 5 most frequently communicated with (measured by estimated # of lines) not actually in top 5 (measured by actual # of lines).

V. Codebooks

There is a distribution tape available with data and data conversion procedures. The raw data file is called NJIT on the tape and is adequately described by the NJIT codebook in Appendix IV. The data can be rearranged in any number of ways for analysis. We have chosen to examine first a subject-by-window, or S x W, approach, where each combination of subject and window makes up one case. (Some variables, like those related to first top 500 hits, have to be brutally forced into the S x W mold.)

The program that accomplishes the S x W conversion is called ACCMS.RUN, and is included on the tape. It produces a file called NJIT.SPSS which consists of the NJIT raw data reorganized into an S x W arrangement. Its format is described in SPSSXW which is an SPSS program (in Appendix V) that reads NJIT.SPSS; provides labels and missing value codes, etc.; produces a codebook; and saves a system file, (already provided if this is an IBM labeled tape) called SXWSAVE. The program file called ACCMS.RUN also computes 48 accuracy measures (hence the name) appropriate for a S x W layout, and stores the result of file SPSS.

The last file on the tape is CLNRT. This simple fortran program reads NJIT and verifies that the data are organized according to the NJIT codebook. It might be a useful place to begin for a programmer who desires to reorganize the data in his or her own way. One warning about the data file called NJIT: we have discovered two incorrect digits on the tape. The program ACCMS.RUN corrects these digits; CLNRT does not.

Accession For	
NTIS GML&I	<input checked="" type="checkbox"/>
DDC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/ _____	
Availability Codes	
Dist	Avail and/or special
A	

Table 1. Summary of Basic Accuracy Measures.

T1	$\frac{T}{F}$ \frac{B}	-- The number of people recalled who were not actually communicated with.
T1P	$\frac{T}{F}$ \frac{B}	-- $T1/NR$, where NR is the number of people recalled. ²
T2	$\frac{T}{F}$ \frac{B}	-- The number of people not recalled who were actually communicated with.
T2P	$\frac{T}{F}$ \frac{B}	-- $T2/NA$, where NA is the number of people actually communicated with.
T12A	$\frac{T}{F}$ \frac{B}	-- $(T1 + T2)/NA$
T12AR	$\frac{T}{F}$ \frac{B}	-- $(T1 + T2)/(NR + NA)$. This represents the percentage of the total possible number of mistakes made by the informant.
TOP5	$\frac{T}{F}$ \frac{B}	<div> <div>M</div> <div>L</div> </div> -- Let a "hit" be defined as a person in both the top 5 most intense recalled and the top 5 most intense actually. Then <div> $TOP5 = 1 - \frac{\text{number of hits}}{5}$ </div>
TOP3	$\frac{T}{F}$ \frac{B}	<div> <div>M</div> <div>L</div> </div> -- Similarly, for top 3, <div> $TOP3 = 1 - \frac{\text{number of hits}}{3}$ </div>
TOP1	$\frac{T}{F}$ \frac{B}	<div> <div>M</div> <div>L</div> </div> -- and <div> $TOP1 = 1 - \frac{\text{number of hits}}{1}$ </div>
WIN2	$\frac{T}{F}$ \frac{B}	<div> <div>M</div> <div>L</div> </div> -- Let a "hit" mean that the rank of a person on the recalled list is within 2 of his or her rank on the actual list. Then <div> $WIN2 = 1 - \frac{\text{number of hits}}{\text{number recalled}}$ </div>
WIN20	$\frac{T}{F}$ \frac{B}	<div> <div>M</div> <div>L</div> </div> -- Let a "hit" mean that the percentile rank of a person on the recalled list is within 10 of his or her rank on the actual list, so that <div> $WIN20 = 1 - \frac{\text{number of hits}}{\text{number recalled}}$ </div>

FOOTNOTES

1. The privacy of the content of those messages is zealously guarded. We do not treat the content of messages in this experiment, only what is known as "who-to-whom traffic," or who communicated with whom, and for how many lines of type.
2. Virtually all of the percentages in this study are what Tukey (1977) calls "started." For example, instead of $TlP = Tl/NR$, we actually use $TlP = (Tl + 1/6)/(NR + 1/3)$. The specific purpose is to make a small adjustment to all of the ratios which will permit later transformation by logs, inverses, ratios, etc., where values of zero cause problems.

APPENDIX I

Letter of Invitation to Participate in Russ' Experiment on EIES

Dear Russ,

I would like to invite you to participate in a social science experiment I am conducting through EIES from January to April, 1979. The experiment is funded by the National Science Foundation. An anthropologist by training, I have been doing social network research over the past few years, and I am a member of EIES Group 35, the Social Network Community. First I'd like to give you a little background information and then describe the experiment for which I'm inviting you to participate.

Most social network data are collected by asking people whom they know, with whom they communicate, with whom they work, and so forth. Then, the structure of the social network can be determined from an analysis of these data. Changes in a social network over time are used to show how effective a particular program or activity was. Furthermore, people working in organizational development have begun to compare a given social network with the organization's "chain of command" charts to provide information about how best to restructure an organization.

There is one major problem with this. In our research, my colleagues and I have found that there is little correspondence between whom people say they talk to over periods of 5 days or more and whom they actually talk to! It's not that people are being devious or trying to keep secrets. Rather, there is simply some as-yet-unmeasured level of forgetting over a period of time. Peter Killworth, Lee Sailer, and I have written a series of papers about our experiments and research to date on this problem. I would be happy to send them to you, even if you decide not to participate in this experiment on EIES.

We would like to measure people's accuracy of recall over different periods of time. This will allow us to calibrate "who do you _____" data collection instruments used by social network researchers. The only way to measure people's accuracy is to have actual communications data for comparison with their recalled communications data. EIES provides the first experimental setting for collecting both kinds of data with ease.

If you consent to volunteer for this experiment, you will be asked to list the names or numbers of all people with whom you communicated by private messages during a given period of time and how many messages you sent to and received from each. You will also be asked how many lines of text you sent to and received from each person you name. There are 28 time periods for which you will be asked, ranging from "during the last two months" to "the last time you were on line." You will also be asked how confident you are in your answers about with whom you communicated, how many messages were sent and received, and how many lines of text were sent and received.

As a way of compensating you for your participation, I will pay you \$50 if you complete the entire experiment. Furthermore, I will be happy to make your own data available to you so you can find out how accurate your recollections are.

Each time you log on EIES, you will be greeted with a short sign-on message such as "no interview this time," or, on a random basis, the automated interview will ask about your communications for a given time period. This will occur before you are asked if you want to list those now on line and before you can accept any waiting communications.

Since we know that it is not always convenient for you to be interviewed, we have made the experiment a little more flexible. First, you will be asked if you want to take a raincheck on the interview for a later time. The procedure will attempt to schedule another interview with you within one week. Only one such raincheck will be allowed per interview session.

Second, we have established the notion of a "harrassment limit." When you give formal consent to participate (through the +PERMIT command described below), you will be asked to set your own harrassment limit for the length of each interview. I would suggest 20 minutes as a reasonable harrassment limit, especially if you are an active communicator. Once an interview starts asking about a given time period (like communication during the last week), you must complete it. At the end of your answers, the procedure will tell you how much of your harrassment limit has been used. If there is time left over for that particular interview session, you will be asked if you want to answer questions about another time period (such as last session on line). Some interviews will take longer than others because of lots of communication you may have had, system slowness, the length of time it takes you to remember your communications, and other such factors beyond our control.

If you volunteer, you will be asked about your communications as described above, and you will also receive two other short questionnaires on line, one at the beginning of the experiment and one at the end. These questionnaires will ask about whether you keep logs of your communications, how well you remember things like birthdays and telephone numbers, your age and sex, and some simple ratings of how important or useful your communication was with each given person.

In order to compare your answers with your actual communications during each given time period, we also need your permission to collect data about your actual communications. If you give such permission, a procedure will look at your message statistics and report exactly the same information we are asking you about: with whom you communicated, how many messages you sent to and received from each person, and how many lines of text you sent to and received from others. The procedure will not look at the keys or the text of any messages, so their content will remain confidential.

To further protect your privacy, any reports using your recall data and your actual communications data will be written so that your identity is not revealed.

To volunteer for this experiment, enter +PERMIT and answer the questions it asks. This is our way of getting your explicit permission to interview you and to collect your actual communications data. If you wish to withdraw from the

experiment completely at some later time, enter +WITHDRAW and your permission will be revoked. If you decide to withdraw, you do so with our thanks for any time you have already spent participating in the experiment. You may also rejoin the experiment by entering +PERMIT once again.

Please feel free to ask me any questions or to discuss this experiment with me further before you volunteer. My EIES number is 357.

I am excited about this unique opportunity to collect data about communications recalled and what actually occurred. These data will help us measure the accuracy of recalled communications over different periods of time.

Sincerely,

H. Russell Bernard (RUSS, 357)

APPENDIX II

Listing of Windows*

WINDOW	HOW LONG	TIME AGO	INTERVIEWS COMPLETED
1	1 MO	2 MOS	36
2	1 MO	1 MO	36
3	2 WKS	2 MOS	35
4	2 WKS	1 MO	35
5	2 WKS	2 WKS	35
6	1 WK	2 MOS	32
7	1 WK	1 MO	34
8	1 WK	2 WKS	35
9	1 WK	1 WK	34
10	3 DYS	2 MOS	34
11	3 DYS	1 MO	36
12	3 DYS	2 WKS	37
13	3 DYS	1 WK	35
14	3 DYS	3 DYS	34
15	2 DYS	2 MOS	36
16	2 DYS	1 MO	36
17	2 DYS	2 WKS	35
18	2 DYS	1 WK	35
19	2 DYS	3 DYS	35
20	2 DYS	2 DYS	34
21	1 DAY	2 MOS	37
22	1 DAY	1 MO	38
23	1 DAY	2 WKS	33
24	1 DAY	1 WK	37
25	1 DAY	3 DYS	33
26	1 DAY	2 DYS	34
27	1 DAY	1 DAY	37
28	LASTON	LASTON	37
29	LASTON	LASTON	34
30	LASTON	LASTON	29
31	LASTON	LASTON	25
32	LASTON	LASTON	25
33	LASTON	LASTON	24
34	LASTON	LASTON	23
35	LASTON	LASTON	23
36	LASTON	LASTON	22
37	LASTON	LASTON	22

*Note: 10 different "last on" windows were asked. The final listing of the experiment shows that 37 persons answered at least one "last on;" 22 persons answered 10 "last ons."

APPENDIX III

An Example of the Use of the "Feedback" Routine

INITIAL CHOICE?+feedback
HERE IS A LIST OF THE #'S OF THE WINDOWS YOU HAVE COMPLETED:
21,6,16,9,4,22,19,1,11,14,27,15,18,20,8,24,7,13,10,12,25,2,3,26,23,17,5,28,29,
30,31,32,33,34,35,36,37
TYPE +WINDOWS FOR A DESCRIPTION OF WHAT THE ABOVE WINDOW #'S MEAN.
#S OF WINDOWS ON WHICH YOU WANT FEEDBACK?5

WINDOW #5 FOR SUBJECT #3 (EIES #357)

DATE OF INTERVIEW: 4/6/79 1:43 PM
FROM: 3/23/79 12:00 AM -- DAY OF WEEK: FRIDAY
TO: 4/5/79 11:59 PM -- DAY OF WEEK: THURSDAY

OF PEOPLE RECALLED: 9
OF ENTITIES ACTUALLY COMMUNICATED WITH: 27
(ENTITY = AN INDIVIDUAL, A GROUP, OR THE CATEGORY OF ANON/PEN/SYSTEM MSGS.)
INTERVIEW STATUS: ACTUAL STATS COLLECTED

CONFIDENCE RATINGS: 2,1,1,1,1
CONNECT MINUTES TOTAL: 12707, MONTH: 87.
INTERVIEW TIME IN SECONDS: 590

RECALLED PRIVATE MESSAGES

EIES #	SENT TO #/LNS	RECEIVED FROM #/LNS
----	-----	-----
114	2/20	2/10
116	2/20	2/30
118	20/200	10/150
213	1/10	0/0
249	2/20	2/20
362	1/10	0/0
369	2/20	2/20
386	3/30	4/30
985	1/10	0/0
TOTALS	34/340	22/260

COMMENTS:

COMMENTS ON WINDOW 5 BY S#3, USER 357.

It told me that 359 was not a valid name....this is probably because robin has left the system with the official 3/31 end date for group 35. But I did communicate with her during the period specified in the interview.

APPENDIX IV

Codebook for NJIT - raw EIES data collected by procedures written by Peter & Trudy Johnson-Lenz. Original resides in EIES N989P202, 208-214. Procedures appear in Appendix V.

CODEBOOK -- BASIC DATA FORMAT

Each card image in the data contains two identification fields: (1) columns 1-2 contain the experimental subject ID number (ranging from 1 through 59), and (2) columns 3-7 contain a within-subject identifier. The within-subject identifiers are defined in the context of card(s) which they identify in the material below. Subject 20 is absent from the data file, since that ID was used for testing purposes and contains no real data. All other IDs are present in the data file. Not all of these completed the entire series of interviews. The status of an ID can be found in columns 8-9 or card 00001 for each subject. N989NP202 contains the list of codes -- a code of 29 represents a subject with complete data. The details of card 00001 are shown below. Other cards are detailed in subsequent pages of this CODEBOOK.

CARD 00001 -- BASIC SUBJECT STATISTICS

<u>columns</u>	<u>contents</u>
1-2	ID# -- for puproses of this experiment
3-7	00001
8-9	status (-2 to 30 -- see N989NP202 for code) on last page
10-11	# of laston interviews completed
12	1=withdrawn from experiment before completion, else 0
13-14	# of rainchecks taken
15-16	# of times used +FEEDBACK
17-18	# of times used +WITHDRAW
19-20	# of times used +HARASS
21-22	# of times used +PERMIT (excluding first time)
23-24	harassment limit, in minutes
25-26	# of times used +WINDOWS
27-29	EIES ID of subject

ACTUAL MESSAGE TRAFFIC

	SENT TO				RECEIVED FROM			
EIES #	PRIVATE #/LNS	P/GRP* #/LNS	P/GRP #/LNS	GROUP #/LNS	PRIVATE #/LNS	P/GRP* #/LNS	P/GRP #/LNS	GROUP #/LNS
104	1/5	0/0	0/0	0/0	0/0	0/0	0/0	0/0
114	3/19	0/0	0/0	0/0	2/9	0/0	0/0	0/0
116	1/7	0/0	0/0	0/0	9/224	0/0	0/0	0/0
118	18/111	0/0	0/0	0/0	18/254	0/0	0/0	0/0
202	0/0	0/0	0/0	0/0	1/5	0/0	0/0	0/0
203	2/9	0/0	0/0	0/0	0/0	0/0	0/0	0/0
213	2/14	0/0	0/0	0/0	0/0	0/0	0/0	0/0
249	3/26	0/0	0/0	0/0	1/19	0/0	0/0	0/0
350	5/49	0/0	0/0	0/0	2/13	0/0	0/0	0/0
352	0/0	0/0	0/0	0/0	1/14	0/0	0/0	0/0
354	1/10	0/0	0/0	0/0	0/0	0/0	0/0	0/0
359	4/86	0/0	0/0	0/0	3/79	0/0	0/0	0/0
362	1/6	0/0	0/0	0/0	0/0	0/0	0/0	0/0
363	1/5	0/0	0/0	0/0	6/62	0/0	0/0	0/0
369	1/12	0/0	0/0	0/0	2/14	0/0	0/0	0/0
370	0/0	0/0	0/0	0/0	1/16	0/0	0/0	0/0
386	5/31	0/0	0/0	0/0	4/58	0/0	0/0	0/0
387	1/5	0/0	0/0	0/0	0/0	0/0	0/0	0/0
389	2/27	0/0	0/0	0/0	0/0	0/0	0/0	0/0
554	1/23	0/0	0/0	0/0	0/0	0/0	0/0	0/0
559	1/23	0/0	0/0	0/0	0/0	0/0	0/0	0/0
612	1/6	0/0	0/0	0/0	0/0	0/0	0/0	0/0
624	2/9	0/0	0/0	0/0	1/1	0/0	0/0	0/0
919	0/0	0/0	0/0	0/0	0/0	1/6	0/0	0/0
983	0/0	0/0	0/0	0/0	2/8	0/0	0/0	0/0
984	0/0	0/0	0/0	0/0	1/4	0/0	0/0	0/0
985	1/7	0/0	0/0	0/0	0/0	0/0	0/0	0/0
TOTALS	57/490	0/0	0/0	0/0	54/780	1/6	0/0	0/0

OF PEOPLE ACTUALLY COMMUNICATED WITH: 27.

(EXCLUDES ANON., PEN, OR SYSTEM MSGS. AND GROUPS TO WHICH YOU SENT MSGS.)

OF DELETED MESSAGES FOUND DURING PROCESSING OF WINDOW: 0.

PRIVATE = PRIVATE MESSAGES: SENT ONLY TO SPECIFIC INDIVIDUALS

P/GRP* = PRIVATE/GRP MSGS.: SENT TO INDIVIDUALS AND GROUPS W/ YOU AS ADDRESSEE.

P/GRP = PRIVATE/GRP MSGS.: SENT TO INDIVIDUALS AND GROUP W/OUT YOU AS ADDRESSEE

GROUP = GROUP MSGS.: SENT ONLY TO GROUP IDS

CARD 00002 -- TIME IN INTERVIEWS -- Part I

Columns contents

1-2	ID#
3-7	00002
8-11	seconds spent in background interview
12-15	seconds spent for window #1
16-19	seconds spent for window #2
20-23	seconds spent for window #3
24-27	seconds spent for window #4
28-31	seconds spent for window #5
32-35	seconds spent for window #6
36-39	seconds spent for window #7
40-43	seconds spent for window #8
44-47	seconds spent for window #9
48-51	seconds spent for window #10
52-55	seconds spent for window #11
56-59	seconds spent for window #12
60-63	seconds spent for window #13
64-67	seconds spent for window #14
68-71	seconds spent for window #15
72-75	seconds spent for window #16
76-79	seconds spent for window #17

CARD 00003 -- TIME IN INTERVIEWS -- PART 2

Columns contents

1-2	ID#
3-7	00003
8-11	seconds spent in window #18
12-15	seconds spent in window #19
16-19	seconds spent in window #20
20-23	seconds spent in window #21
24-27	seconds spent in window #22
28-31	seconds spent in window #23
32-35	seconds spent in window #24
36-39	seconds spent in window #25
40-43	seconds spent in window #26
44-47	seconds spent in window #27
48-51	seconds spent in window #28
52-55	seconds spent in window #29
56-59	seconds spent in window #30
60-63	seconds spent in window #31
64-67	seconds spent in window #32
68-71	seconds spent in window #33
72-75	seconds spent in window #34
76-79	seconds spent in window #35

CARD 00004 -- TIME IN INTERVIEWS -- PART 3

Columns contents

1-2 ID#
 3-7 00004
 8-11 seconds spent in window #36
 12-15 seconds spent in window #37
 16-19 seconds spent in follow-up interview

CARD 00005 -- BACKGROUND INTERVIEW

Columns contents

1-2 ID#
 3-7 00005
 8 how well remember ZIPCODES (1=not very well, to 7=very well)
 9 how well remember PHONE #s (1=not very well, to 7=very well)
 10 how well remember NAMES (1=not very well, to 7=very well)
 11 how well remember FACES (1=not verywell, to 7=very well)
 12 how well remember DATES (1=not very well, to 7=very well)
 13 how well remember LYRICS (1=not very well, to 7=very well)
 14 how well remember BIRTHDAYS, ETC. (1=not very well, to 7=very well)
 15 keep logs of EIES communications? (1=yes, 0=no)
 16 keep files of EIES communications? (1=yes, 0=no)
 17 use a CRT for EIES work? (1=yes, 0=no)
 18 sex (1=male, 2=female, 0=refused to answer)
 19-20 age in years (0=refused to answer)

CARD 00006 -- FOLLOWUP INTERVIEW -- Part 1

Columns contents

1-2 ID#
 3-7 00006
 8 log EIES communications differently? (1=yes, 0=no)
 9 file EIES communications differently? (1=yes, 0=no)
 10-12 EIES ID of #1 of top 20
 13-15 EIES ID of #2 of top 20
 16-18 EIES ID of #3 of top 20
 19-21 EIES ID of #4 of top 20
 22-24 EIES ID of #5 of top 20
 25-27 EIES ID of #6 of top 20
 28-30 EIES ID of #7 of top 20
 31-33 EIES ID of #8 of top 20
 34-36 EIES ID of #9 of top 20
 37-39 EIES ID of #10 of top 20
 40-42 EIES ID of #11 of top 20
 43-45 EIES ID of #12 of top 20
 46-48 EIES ID of #13 of top 20
 49-51 EIES ID of #14 of top 20
 52-54 EIES ID of #15 of top 20
 55-57 EIES ID of #16 of top 20
 58-60 EIES ID of #17 of top 20

note -- the Top 20 are those with whom
 the subject communicated most during
 the experiment.

61-63 EIES ID of #18 of top 20
 64-66 EIES ID of #19 of top 20
 67-69 EIES ID of #20 of top 20

(note -- there is no particular order of the top 20 IDs -- due to variations in processing from subject to subject no order is implied. Furthermore, some subjects did not communicate with 20 people, hence fewer than 20 IDs are shown. Remaining unused portions are filled with blanks.)

CARD 00007 -- FOLLOWUP INTERVIEW -- PART 2

Columns contents

1-2 ID#
 3-7 00007
 8-11 ratings of #1 of top 20

(first rating, column 8, is rating 1=unimportant to 7=important)
 (second rating, column 9, is 1=unsatisfying to 7=satisfying)
 (third rating, column 10, is 1=undesirable to 7=desirable)
 (fourth rating, column 11, is 1=uninteresting to 7=interesting)

12-15 ratings of #2 of top 20 (same scales as above)
 16-19 ratings of #3 of top 20 "
 20-23 ratings of #4 of top 20 "
 24-27 ratings of #5 of top 20 "
 28-31 ratings of #6 of top 20 "
 32-35 ratings of #7 of top 20 "
 36-39 ratings of #8 of top 20 "
 40-43 ratings of #9 of top 20 "

44-47 ratings of #10 of top 20 (same scales as above)
 48-51 ratings of #11 of top 20 "
 52-55 ratings of #12 of top 20 "
 56-59 ratings of #13 of top 20 "
 60-63 ratings of #14 of top 20 "
 64-67 ratings of #15 of top 20 "
 68-71 ratings of #16 of top 20 "
 72-75 ratings of #17 of top 20 "
 76-79 ratings of #18 of top 20 "

CARD 00008 -- FOLLOWUP INTERVIEW -- PART 3

Columns contents

1-2 ID#
 3-7 00008
 8-11 ratings of #19 of top 20 (same scales as above)
 12-15 ratings of #20 of top 20 "
 16-35 date/time of interview
 36-41 Julian days of date of interview

CARD 00009 -- ORDER OF WINDOW PRESENTATION

Columns contents

1-2 ID#
 3-7 00009
 8-9 index of first window presented
 10-11 index of second window presented
 *** ***
 60-61 index of 27th window presented

CARDS XX000 -- WINDOW INTERVIEW -- BASIC INFORMATION

Columns contents

1-2 ID#
 3-7 XX000 (where XX is window # 1-37)
 8 confidence in list of people (1=not at all, 7=very)
 9 confidence in messages sent (1=not at all, 7=very)
 10 confidence in messages received (1=not at all, 7=very)
 11 confidence in lines sent (1=not at all, 7=very)
 12 confidence in lines received (1=not at all, 7=very)
 13-15 # of cognitive communicants
 16-18 # of behavioral communicants (incl. anon + GRPS sent to)
 19 status (0=cognitive in progress, 1=cognitive complete
 2=behavioral complete, 3=no behavioral - s. withdrawn
 4=behavioral in progress)
 20-21 # of deleted messages found during processing window
 22-27 connect minutes used to date of interview, i.e., EIES experience
 28-33 connect minutes used to date in current month
 34-53 date/time of interview
 54-59 Julian days of date of interview

CARD XX001 -- WINDOW INTERVIEW -- DATES OF WINDOW

Columns contents

1-2 ID#
 3-7 XX001 (where XX is window #)
 8-27 date/time of start of window
 28-47 date/time of stop of window
 48-53 Julian days of date of start of window
 54-59 Julian days of date of stop of window
 60 day of week of start of window (1=Sunday, 7=Saturday)
 61 day of week of stop of window (1=Sunday, 7=Saturday)

CARDS XX1## -- WINDOW INTERVIEW -- COGNITIVE DATA

Columns contents

1-2 ID#
 3-7 XX1## (where XX is window # and ## is sequential index)
 (XX and ## have leading blanks if only 1 digit)
 8-10 EIES ID of communicant
 11-13 recalled # of messages sent
 14-16 recalled # of messages received
 17-21 recalled # of lines sent
 22-26 recalled # of lines received

CARDS XX2## -- WINDOW INTERVIEW -- BEHAVIORAL DATA

Columns contents

1-2 ID#
 3-7 XX2## (where XX is window # and ## is sequential index)
 (XX and ## have leading blanks if only 1 digit)
 8-10 EIES ID of communicant
 11-13 # of private messages sent
 14-16 # of private messages received P/GRP -- Message with both
 17-21 # of private lines sent private and group addressees
 22-26 # of private lines received
 27-29 # of p/grp* messages sent
 30-32 # of p/grp* messages received
 33-37 # of p/grp* lines sent * --
 38-42 # of p/grp* lines received
 43-45 # of p/grp messages sent
 46-48 # of p/grp messages received
 49-53 # of p/grp lines sent
 54-58 # of p/grp lines received
 59-61 # of group messages sent
 62-64 # of group messages received GROUP -- Message with no
 65-69 # of group lines sent private addressee
 70-74 # of group lines received

SUBJECT STATUS CODES

CODE	STATUS
-2	NOT PERMITTED
-1	ONLY PERMITTED
0	BACKGROUND DONE
1	FIRST REGULAR WINDOW DONE
2	SECOND REGULAR DONE....
27	27TH REGULAR WINDOW DONE
28	ALL 10 LAST ONS DONE IN ADDITION TO REGULAR 27
29	FOLLOW-UP DONE
30	WITHDRAWN

This SPSS program creates a system save file from SPSS, and incidentally provides a description of the variables.

```
//LDSSXW JOB (SC020091,****),'S A I L E R',REGION=256K,TIME=5
/*AFTER LDSCNVT
/*JOBPARM I=5,L=5
/*ROUTE XEQ DISK00
/*ROUTE PRINT RMT99
/*ROUTE PRINT HOLD
/*ROUTE XEQ 2-3420
// EXEC SPSS
//FT08F001 DD DSN=SPSS,VOL=SER=LB2670,UNIT=3400-6,LABEL=(2,SL),
// DISP=(OLD,KEEP)
//FT04F001 DD DSN=SXWSAVE,VOL=SER=KM2697,UNIT=3400-6,
// DISP=(NEW,KEEP),LABEL=(2,SL)
//GO.SYSIN DD *
RUN NAME          MAKE CODEBOOK
FILE NAME         XPT
VARIABLE LIST     EGO STAT1 LASTON WTHD RAINN FEEDN WTHDN HRSSN PERMN
                  HRSSLM WINDN EIPSN BKTIM FOLTIM WINTIM ZIP PHN
                  NAMES FACES DATES LYRICS BRTHDS LOG FILE CRT SEX
                  AGE LOGDIF FILDIF TOP20 IMPORT SATIS DESBLE INTRST
                  FOLDAT ORDER WIN CONL COMMS CONMR CONLS CONLR
                  COGN ACTN STAT2 DLEN CNNCT CNMTH INTDAT  START STOP
                  DAYST
                  DAYSTP T1T T1TP T2T T2TP T12AT T12ART
                  TOP1TM TOP3TM TOP5TM W2TM W20TM
                  TOP1TL TOP3TL TOP5TL W2TL W20TL
                  T1F T1FP T2F T2FP T12AF T12ARF
                  TOP1FM TOP3FM TOP5FM W2FM W20FM
                  TOP1FL TOP3FL TOP5FL W2FL W20FL
                  T1B T1BP T2B T2BP T12AB T12ARB
                  TOP1BM TOP3BM TOP5BM W2BM W20BM
                  TOP1BL TOP3BL TOP5BL W2BL W20BL

INPUT MEDIUM     TAPE
N OF CASES       UNKNOWN
INPUT FORMAT      FIXED(F3.0,6X,2F2.0,F1.0,7F2.0,F3.0/
                  T10,3F4.0,11F1.0,F2.0/
                  T10,2F1.0,F4.0,4F3.1,F6.0,F3.0/
                  T10,F3.0,5F1.0,2F3.0,F1.0,F2.0,5F6.0,2F1.0,
                  3(/F5.0,F5.2,F5.0,13F5.2) )

MISSING VALUES  FOLTIM SEX AGE IMPORT SATIS DESBLE INTRST FOLDAT
                  ORDER DAYST DAYSTP (0)/ T1T TO W20BL (-.01,-.5,-1.0)/
                  TOP20 (-100.0)

VALUE LABELS     STAT1 (-2)NOT PERM (-1)ONLY PERM (0) BCKDON
                  (28)LASTONS (29)FOLLOW-UP (30)WITHDRAWN/
                  WTHD (1) YES (0) NO/ORDER (0) LASTONS/
                  ZIP TO BRTHDS (7) VERY WELL (1) NOT VERY WELL/
                  SEX (1)MALE(2)FEMALE(0)REFUSED/
                  LOG FILE CRT LOGDIF FILDIF (1)YES (0)NO/
                  AGE (0)REFUSED/
```

References Cited

- Bernard, H. Russell and P. D. Killworth
1977 Informant accuracy in social networks data II. Human Communication Research, 4:3-18.
1979 Informant accuracy in social network data IV. Journal of Social Networks, 2 (Winter), in press.
- Killworth, Peter D. and H. R. Bernard
1976 Informant accuracy in social network data. Human Organization, 35:269-286.
1979 Informant accuracy in social network data III. Journal of Social Networks, 2 (Summer), in press.
- Starr, Roxanne Hiltz and Murray Turoff
1978 The Network Nation. Reading, Mass.: Addison-Wesley.
- Tukey, John W.
1977 Exploratory Data Analysis. Reading, Mass.: Addison Wesley.

	IMPORT MEAN IMPORTANCE OF THOSE IN TOP20/
	SATIS MEAN SATISFACTION WITH THOSE IN TOP20/
	DESBLE MEAN DESIRABILITY OF THOSE IN TOP20/
	INTRST MEAN INTEREST IN THOSE IN TOP20/
	CNNCT SUBJECT'S TOTAL EIES CONNECT TIME IN MINUTES/
	CNNMTH SUBJECT'S CONNECT TIME THIS MONTH, SO FAR/
	INTDAT JULIAN DATE OF INTERVIEW FOR THIS WINDOW/
	START JULIAN DATE OF START OF WINDOW/
	STOP JULIAN DATE OF END OF WINDOW/
TASK NAME	SOME VARIABLES' FREQUENCIES
FREQUENCIES	GENERAL=EGO TO EIESN ZIP TO FILDIF FOLDAT TO DLTN
	DAYST TO DAYSTP
OPTIONS	3,6
STATISTICS	1,3,5
READ INPUT DATA	
CONDESCRIPTIVE	BKTIM FOLTIM WINTIM TOP20 TO INTRST CNNCT TO STOP
	T1T TO W20BL
STATISTICS	ALL
SAVE FILE	
FINISH	
&	

VAR LABELS

CONL TO CONLR (7) VERY (1) NOT AT ALL/
 STAT2 (0) COGNITIVE IN PROGRESS (2) COGNITIVE COMPLETE
 (2) BEHAVIORAL COMPLETE (3) WITHDRAWN
 (4) BEHAVIORAL IN PROGRESS /
 DAYST DAYSTP (1) SUNDAY (2) MONDAY (3) TUESDAY
 (4) WEDNESDAY (5) THURSDAY (6) FRIDAY (7) SATURDAY/
 EGO SUBJECT NUMBER/ STAT1 XPT PROGRESS/
 LASTON # OF LAST ON WINDOWS COMPLETED/
 WTHD WITHDREW BEFORE COMPLETION/ RAINN # OF RAINCHECKS/
 FLEDN NUMBER OF TIMES FEEDBACK USED/
 WTHDN NUMBER OF TIMES WITHDRAW USED/
 HRSSN NUMBER OF TIMES HARASS USED/
 PERMN NUMBER OF TIMES PERMIT USED/
 HRSSLM HARASSMENT LIMIT, IN MINUTES/
 WINDN NUMBER OF TIMES WINDOW USED/
 EIESN EIES USER ID#/
 ZIP ABILITY TO RECALL ZIP NUMBERS/
 PHN ABILITY TO RECALL PHONE NUMBERS/
 NAMES ABILITY TO RECALL NAMES/
 FACES ABILITY TO RECALL FACES/
 DATES ABILITY TO RECALL DATES/
 LYRICS ABILITY TO RECALL LYRICS/
 BRTHDS ABILITY TO RECALL BIRTHDAYS, ETC./
 LOG DO YOU LOG YOUR MESSAGES?/
 FILE DO YOU FILE YOUR MESSAGES?/
 CRT DO YOU USUALLY USE A CRT TERMINAL?/
 SEX USUAL PHYSICAL GENDER OF SUBJECT/
 AGE AGE OF SUBJECT/
 LOGDIF DO YOU LOG MESSAGES DIFFERENTLY NOW?/
 FILDIF DO YOU FILE MESSAGES DIFFERENTLY NOW?/
 FOLDAT JULIAN DATE OF FOLLOW-UP INTERVIEW/
 ORDER PRESENTATION SEQUENCE OF THIS WINDOW/
 WIN WINDOW NUMBER OF THIS WINDOW/
 CONL CONFIDENCE IN LIST OF COMMUNICANTS/
 CONMS CONFIDENCE IN NUMBER OF MESSAGES SENT/
 CONMR CONFIDENCE IN NUMBER OF MESSAGES RECEIVED/
 CONLS CONFIDENCE IN NUMBER OF LINES SENT/
 CONLR CONFIDENCE IN NUMBER OF LINES RECEIVED/
 COGN NUMBER OF COMMUNICANTS RECALLED/
 ACTN NUMBER OF ACTUAL COMMUNICANTS/
 STAT2 PROGRESS OF DATA COLLECTION FOR THIS WINDOW/
 DLTN NUMBER OF MESSAGES DELETED IN THIS WINDOW/
 DAYST DAY OF WEEK WINDOW STARTS ON/
 DAYSTP DAY OF WEEK WINDOW STOPS ON/
 BKTIM NUMBER OF SECONDS FOR BACKGROUND INTERVIEW/
 FOLTIM NUMBER OF SECONDS FOR FOLLOW-UP INTERVIEW/
 WINTIM NUMBER OF SECONDS TO TAKE THIS WINDOW/
 TOP20 % OF ACTUAL COMMUNICANTS IN TOP20 LIST/